



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Separating Content and Structure in Humor Appreciation

Heintz, Sonja

Abstract: The present study empirically tests a bimodal model of humor appreciation. In this model, individual differences in evaluating jokes and cartoons are attributed to their contents (sexual and aggressive) and structures (incongruity-resolution and nonsense). A total of 225 participants (64% women, Mage = 32.37 years) rated 50 jokes and cartoons on funniness, offensiveness, and boringness. They also completed a standard humor appreciation measure (the 3 WD). Using structural equation modeling, the bimodal model was found to be superior to alternative models. Regressions predicting the 3 WD categories supported the convergent and discriminant validity of the content and structure factors. In line with previous research, gender differences were found for the content factors, but not for the structure factors. Overall, this study is the first to show the viability of bimodal models of humor appreciation. They allow capturing the sources of individual differences in humor appreciation more adequately than previous models, thus providing a basis for future theories, research, and applications.

DOI: <https://doi.org/10.1027/1614-0001/a000301>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-174382>

Journal Article

Accepted Version

Originally published at:

Heintz, Sonja (2020). Separating Content and Structure in Humor Appreciation. *Journal of Individual Differences*, 41(1):37-44.

DOI: <https://doi.org/10.1027/1614-0001/a000301>

Separating content and structure in humor appreciation: A bimodal structural equation
modeling approach

Sonja Heintz

Department of Psychology, University of Zurich, Zurich, Switzerland

Author Note:

Sonja Heintz is at the Department of Psychology, University of Zurich, Switzerland, Binzmühlestrasse 14/7, 8050 Zurich, Switzerland.

Correspondence concerning this article should be addressed to: Sonja Heintz, Section on Personality and Assessment, Department of Psychology, University of Zurich, Binzmühlestrasse 14/7, 8050 Zurich, Switzerland. E-mail: s.heintz@psychologie.uzh.ch.

Acknowledgements:

I would like to thank Willibald Ruch for his helpful feedback on the design of study, Alex Junghans for his help in collecting and selecting the humorous stimuli, and the raters for their help in selecting the final set of humorous stimuli.

The manuscript was published in the *Journal of Individual Differences*, and the final manuscript version can be obtained at <https://doi.org/10.1027/1614-0001/a000301>

Abstract

The present study empirically tests a bimodal model of humor appreciation. In this model, individual differences in evaluating jokes and cartoons are attributed to their contents (sexual and aggressive) and structures (incongruity-resolution and nonsense). A total of 225 participants (64% women, mean age 32.37 years) rated 50 jokes and cartoons on funniness, offensiveness, and boringness. They also completed a standard humor appreciation measure (the 3 WD). Using structural equation modeling, the bimodal model was found to be superior to alternative models. Regressions predicting the 3 WD categories supported the convergent and discriminant validity of the content and structure factors. In line with previous research, gender differences were found for the content factors, but not for the structure factors. Overall, this study is the first to show the viability of bimodal models of humor appreciation. They allow capturing the sources of individual differences in humor appreciation more adequately than previous models, thus providing a basis for future theories, research, and applications.

Keywords: humor appreciation; individual differences; structural equation modeling; construct validity; behavior tests

Separating content and structure in humor appreciation: A bimodal structural equation modeling approach

Humor is a ubiquitous everyday phenomenon, which broadly entails everything funny (see Martin & Ford, 2018; Ruch, 2007, 2012). Individual differences in humor can be studied in different domains, including humor comprehension (how people understand humor), humor appreciation (how people evaluate humorous stimuli and events), and humor production (how people show humor). The present study focuses on humor appreciation, which has long been of interest to personality researchers (e.g., Cattell & Tollefson, 1966; Eysenck, 1942, 1944; Ruch, 1992; Ruch & Hehl, 2007). This research has shown that reactions to humorous stimuli (e.g., funniness ratings of jokes) are stable individual differences that relate to other personality traits, attitudes, intelligence, and art preferences. They can thus serve as indirect (or in Cattell's sense objective) measures of personality (Cattell & Tollefson, 1966). Furthermore, the properties of humorous stimuli (mainly content and structure) were found to influence these individual differences. The present study extends this research by empirically testing a bimodal model of humor appreciation (as proposed by Ruch & Hehl, 2007; Ruch & Platt, 2012). In this bimodal model, both contents (sexual and aggressive) and structures (incongruity-resolution and nonsense) of humorous stimuli are considered simultaneously.

Content and Structure in Humor Appreciation

The content and structure of humorous stimuli represent two sources that account for a large amount of variance in individual differences in humor appreciation (Carretero-Dios, Pérez, & Buéla-Casal, 2010; Ruch, 1992; Ruch & Hehl, 2007). The content of humorous stimuli is an affective stimulus property with many variations; for example, a stimulus can be

innocent/harmless, benevolent, sexual, satirical, sick, dark, scatological, or disparaging/aggressive/hostile. The content is the most salient stimulus property and has been incorporated in early classifications (e.g., Mindess, Miller, Turek, Bender, & Corbin, 1985) and humor theories (for an overview, see Ferguson & Ford, 2008). Empirically, however, the structure of humorous stimuli was found to be more relevant for explaining individual differences (Ruch, 1992; Ruch & Hehl, 2007). Structure is a cognitive stimulus property and refers to how the joke “works” (i.e., the mechanisms underlying the punchline). This includes linguistic techniques (e.g., aggregation, repetition, and contradiction), incongruities (i.e., punchlines that are surprising and unexpected), and incongruity-resolution (i.e., making sense of punchlines). These different mechanisms were formalized in detail in several models and theories of humor appreciation (e.g., Freud, 1905; Raskin, 1985; Suls, 1983).

A similar distinction of cognitive (structure) and affective properties (contents) was also found in ratings of humorous materials. For example, Wicker et al. (1981) had two sets of jokes rated on 13 scales. They found one cognitive response dimension (originality, which included funniness) and two affective dimensions (emotionality and superiority). Later, Ruch and Rath (1993) had jokes and cartoons rated on 17 scales. They found one positive reaction to humor (funniness), which combined both affective and cognitive elements, and two negative reactions, which were either cognitive (boringness) or affective (offensiveness). It is thus important to go beyond “funniness” ratings to adequately capture different reactions to humorous stimuli.

Integrating Content and Structure

Different theoretical and empirical approaches aimed at integrating both contents and structures (e.g., Carretero-Dios et al., 2010; Eysenck, 1942; Freud, 1905; Godkewitsch, 1976;

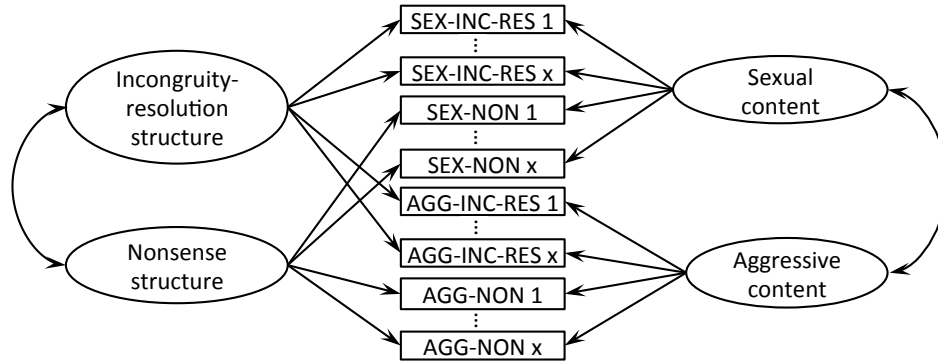
Ruch, 1992; Wicker, Thorelli, Barron, & Ponder, 1981). These approaches often derived content and structure dimensions based on principal component or exploratory factor analyses of ratings of humorous stimuli. Ruch (1992) had large sets of jokes and cartoons rated for funniness and aversiveness, resulting in the 3 WD (3 Witz-Dimensionen) humor test. He distinguished three stimulus categories: Incongruity-resolution (INC-RES), which completely resolves incongruities, nonsense (NON), which contains partially or completely unresolved incongruities, and sexual humor (SEX), which contains sexual contents.

Extending this approach, Carretero-Dios et al. (2010) added the three content categories of black, man-disparagement, and woman-disparagement humor. Thus, two structures (incongruity-resolution and nonsense) and several contents of humorous stimuli (mostly related to sex and aggression; see also Ferguson, & Ford, 2008; Freud, 1905; Godkewitsch, 1976) were found to influence individual differences in humor appreciation.

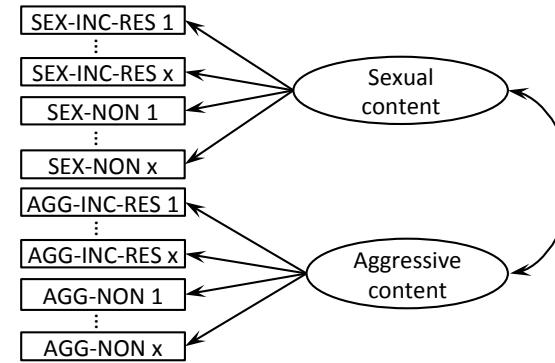
Extending these approaches, it was suggested that content and structure can contribute simultaneously to the appreciation of humorous stimuli with varying significance or salience (Eysenck, 1942; Freud, 1905; Godkewitsch, 1976; Ruch & Hehl, 2007; Ruch & Platt, 2012). However, existing classifications and measures usually only consider one of these properties in each stimulus; that is, a joke that has an incongruity-resolution structure and sexual content might be assigned to the SEX category because the content might be more dominant; conversely, a cartoon with a nonsense structure and aggressive content might be assigned to the NON category because the structure is more dominant. Modeling both the structure and content properties of these stimuli simultaneously allows a more adequate representation and assessment of the sources of variance contributing to humor appreciation.

The Present Study

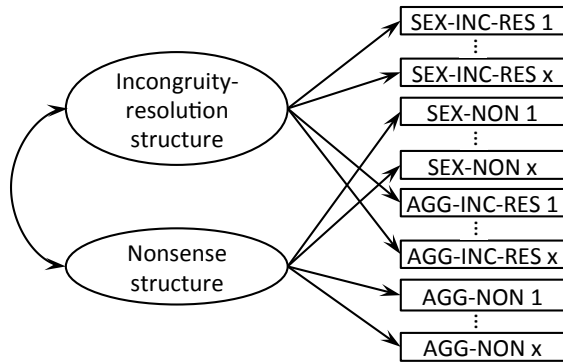
A bimodal model of humor appreciation is needed to assess the contributions of both content and structure to individual differences in humor appreciation in each stimulus (see Ruch & Hehl, 2007; Ruch & Platt, 2012). This allows a better understanding of what makes humor funny, offensive, and boring to different people. This knowledge can then be used in future theories (e.g., models of humor appreciation), research (e.g., classifications of humor stimuli) and applications (e.g., selecting appropriate stimuli for humor interventions). The present study is the first to investigate whether such a bimodal model of humor appreciation can be empirically supported. Specifically, a set of jokes and cartoons that contain two contents (SEX and aggression, AGG) and two structures (INC-RES and NON) are rated for funniness, offensiveness, and boringness. These contents and structures were selected as the most prevalent ones in theoretical and empirical work on humor appreciation. To test the model, four different structural equation models are compared separate for each of the three ratings: Structure-only, content-only, four-factor, and bimodal models (see Figure 1). It is expected that the bimodal model will fit the data best.



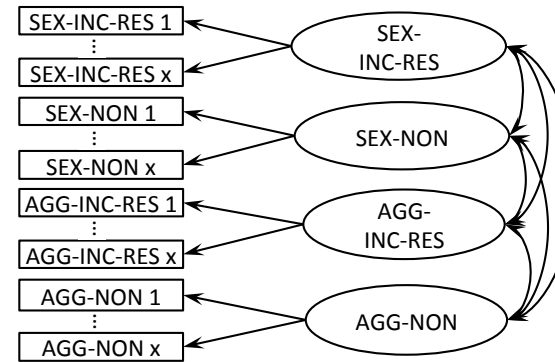
(a) Bimodal model with two contents and two structures



(b) Unimodal model (content-only)



(c) Unimodal model (structure-only)



(d) Four-factor model

Figure 1. Visualization of the four competing SEM models tested in the present study. (a) Bimodal model with two contents and two structures; (b) Unimodal model (content-only); (c) Unimodal model (structure-only); (d) Four-factor model.

To test the convergent and discriminant validity of this bimodal model of humor appreciation, the two content and structure factors are related to the 3 WD as a standard test of humor appreciation (Ruch, 1992). It is expected that the largest correlations occur between the funniness ratings of the INC-RES, NON, and SEX factors and categories, respectively. The relationship between the offensiveness and boringness ratings in the bimodal model and the aversiveness ratings in the 3 WD are investigated exploratorily, though the largest correlations are again expected among the corresponding factors and categories (INC-RES, NON, and SEX). The relationship of the AGG factor with the 3 WD categories is expected to be lower, as the 3 WD contains only few aggressive contents that are spread across all three categories.

Finally, gender differences in humor appreciation are explored. A recent systematic review showed that men and women appreciated INC-RES and NON humor to a similar extent, while men appreciated SEX and AGG humor more than women did (Hofmann, Platt, Lau, & Torres-Marín, 2019). Gender differences are thus expected for the two content factors (with men appreciating them more than women), but not for the two structure factors.

Methods

Sample

A total of 232 participants completed the humorous stimuli. Seven participants were excluded from further analyses because they a) provided implausible responses ($n=1$), b) completed the study too quickly ($n=1$), or c) indicated an age of less than 18 years ($n=5$). The final sample thus consisted of 225 participants (64% women and 36% men) with a mean age of 32.37 years ($SD=12.20$, range 18–87 years). Most participants were Swiss (50%) or German (43%).

They were mostly well educated, with 45% having a university degree, 23% currently studying, 20% having a university-entrance diploma, and 9% having an apprenticeship.

Measures

Humorous stimuli. The stimuli were selected in a two-step procedure. In the first step, two raters judged a large pool of jokes and cartoons regarding their content and structure to ensure the content validity of the items (see the supplementary material for a detailed description of the initial stimuli selection). Of the final set of 50 jokes and cartoons, 13 contained sexual and nonsense humor (SEX-NON; example joke “How do you recognize a gay snowman? He has the carrot in his buttocks.”), 15 contained aggressive and nonsense humor (AGG-NON; e.g., “Three snails on a rail track. The first one says ‘Take care, a train is coming’ Crack! – ‘Where?’ Crack! – ‘There!’ Crack!”), 11 contained sexual and incongruity-resolution humor (SEX-INC-RES; e.g., “As her partner was getting sexually exhausted, the women threw a coin out of the window to the snake charmer and yelled ‘Continue playing a bit longer’”), and 11 contained aggressive and incongruity-resolution humor (AGG-INC-RES; e.g., “The leading candidate of the party is speaking. His advertising consultant whispers: ‘Yesterday was better.’ – ‘But I didn’t give a speech yesterday!’ – ‘Exactly!’”). Each stimulus was rated for funniness, offensiveness, and boringness (see Ruch & Rath, 1993) on five-point scales (from 1=*not at all* to 5=*very much*). Four additional stimuli (one from each category) were employed as a warm-up and were not scored. The Cronbach’s alphas, means, standard deviations, and intercorrelations among the initial 12 content×structure× rating combinations are shown in Supplementary Table S1. The internal consistencies were sufficient for all combinations (.73–.94, *Mdn*=.85), yet the correlations

among the combinations of the same ratings were high as well (all r s .54–.89, all p s<.01, Mdn =.70).

In a second step, these 50 stimuli were empirically refined by selecting the ones that were most representative for each combination and that differed from the other combinations (see the supplementary material for a detailed description of the empirical stimuli selection; see Reiss & Reips, 2016). This resulted in a total of 29 stimuli for the funniness ratings (4–9 for each combination), 33 for offensiveness (4–11 each), and 24 for boringness (3–8 each). The Cronbach's alphas, means, standard deviations, and intercorrelations among the 12 refined content×structure×rating combinations are shown in Supplementary Table S2. Again, the internal consistencies were sufficient for all combinations (.59–.93, Mdn =.74), while the correlations among the combinations of the same ratings were reduced (all r s .32–.73, all p s<.01, Mdn =.55). Thus, a median 11% loss of true-score variance (which is expected due to the lower item number) was compensated by a reduction of the median overlap between the combinations of 19% (i.e., a better differentiation between the content×structure combinations). Furthermore, confirmatory factor analyses supported that the resulting stimuli are unidimensional (see Supplementary Table S3).

3 WD (Ruch, 1992). The 3 WD assesses the appreciation of jokes and cartoons of the three humor categories INC-RES, NON, and SEX. Ten cartoons and jokes for each dimension plus five warm-up stimuli (not scored) are rated for funniness and aversiveness using seven-point Likert-type scales (from 1 to 7). Three funniness and three aversiveness scores were computed (for INC-RES, NON, and SEX). Cronbach's alpha ranged from .79 (NON funniness) to .92 (INC-RES aversiveness), and the means ranged from 1.68 (INC-RES aversiveness, SD =1.00) to 3.50 (INC-RES funniness, SD =1.33).

Procedure

The study was conducted online (www.unipark.info). Participants first completed the demographic items, followed by the 54 humorous stimuli. Each stimulus was shown on a separate page, followed by the three ratings (funniness, offensiveness, and boringness). Afterwards, participants completed the 3 WD, also with each stimulus shown on a separate page. Other variables were assessed that are not relevant for the present study. Participants could receive a general feedback on the results of the study and course credit (for psychology students). An a-priori power analysis was not feasible to determine the sample size, but at least 200 participants were recruited to conduct the structural equation modeling (SEM) analyses. The study was conducted in line with the local ethical guidelines.

Analyses

The bimodal, unimodal (content-only and structure-only), and four-factor models (one for each combination) were computed separately for the funniness, offensiveness, and boringness ratings. The two latent content (SEX and AGG) and structure factors (INC-RES and NON) were allowed to correlate with one another, respectively, but correlations between the content and structure factors were set to zero to ensure model identification. The convergent and discriminant validity with the 3 WD categories as well as the criterion validity with gender were tested in SEM regression analyses. Each of these variables were added as observed criteria in the bimodal SEM, and the four latent factors (SEX, AGG, INC-RES, and NON) served as predictors.

The robust maximum likelihood estimator was used in all SEMs to accommodate deviations from the normal distribution (especially in the offensiveness ratings). The analyses were conducted in *R* (R Core Team, 2018), using the packages *psych* (Revelle, 2018) and

lavaan (Rosseel, 2012). Fit indices were interpreted in line with the guidelines of Schermelleh-Engel, Moosbrugger, and Müller (2003): χ^2 (good: $p > .05$, acceptable: $p \geq .01$), χ^2/df (good: ≤ 2 , acceptable: ≤ 3), comparative fit index (CFI; good: $\geq .97$, acceptable: $\geq .95$), root mean square error of approximation (RMSEA; good: $\leq .05$, acceptable: $\leq .08$), and standardized root mean square residual (SRMR; good: $\leq .05$, acceptable: $\leq .10$). The BIC (sample-size adjusted Bayesian information criterion) was also computed, for which smaller values indicate a better fit to the data.

Results

Model comparisons

Table 1 shows the model fits of the unimodal (content-only or structure-only), bimodal, and four-factor models of the three ratings. For all ratings (funniness, offensiveness, and boringness), the bimodal models showed an acceptable fit. The two structures (INC-RES and NON) correlated positively ($r = .42-.59$, all $ps < .01$) and shared 18–35% of their true-score variance. The contents ($r = .66-.75$, all $ps < .01$) were also positively correlated and shared 44–56% of their true-score variance. These overlaps might represent general humor appreciation, as some people (e.g., those higher in cheerfulness; Ruch & Hofmann, 2012) tend to enjoy humorous stimuli more than others, independent of its content and structure.

Alternative unimodal models that specified only two content factors or only two structure factors and a model that specified four latent factors (one for each content×structure combination) showed a worse fit to the data, as indexed by all absolute and relative fit indices. These results support the hypothesis that both contents and structures should be distinguished simultaneously. A bimodal model adequately represented the variance in humor appreciation due to the content and structure properties within each humorous stimulus.

Table 1

Model Fits of the Bimodal, Unimodal (Content-only or Structure-only), and Four-Factor-Models of the Three Ratings (Funniness, Offensiveness, and Boringness)

Rating	Model	χ^2	df	χ^2/df	CFI	RMSEA	SRMR	BIC
Funniness	Bimodal	533.80***	346	1.54	.847	.049	.068	18,495
Funniness	Structure	674.98***	376	1.80	.757	.059	.064	18,593
Funniness	Content	766.48***	376	2.04	.682	.068	.086	18,710
Funniness	4 factors	617.40***	371	1.66	.800	.054	.076	18,531
Offensiveness	Bimodal	564.19**	460	1.23	.920	.032	.060	15,401
Offensiveness	Structure	755.30***	494	1.53	.798	.048	.095	15,776
Offensiveness	Content	676.51***	494	1.37	.859	.041	.082	15,563
Offensiveness	4 factors	656.11***	489	1.34	.871	.039	.077	15,528
Boringness	Bimodal	320.99***	226	1.42	.945	.043	.053	17,481
Boringness	Structure	431.53***	251	1.72	.895	.057	.066	17,559
Boringness	Content	532.50***	251	2.12	.837	.071	.076	17,665
Boringness	4 factors	390.64***	246	1.59	.916	.051	.064	17,518

Note. $N=225$. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; BIC = sample-size adjusted Bayesian information criterion.

** $p < .01$. *** $p < .001$.

Furthermore, some stimuli were mostly loaded by the content and others by the structure factor, which also varied between the three ratings (see Supplementary Tables S4–S6 for details). For example, Stimulus 27 (“Dad asks Max: ‘What did you do in physics today?’–‘Built a bomb!’–Dad: ‘Alright, and what will you do tomorrow at school?’–Max: ‘Which school?’”) was found funny both because of its structure (INC-RES) and its content (AGG), while people were offended by it mostly due its content, and they were bored by it mostly due its structure. To explore this pattern across the ratings, an arbitrary cut-off of $\geq .10$ was set to interpret loadings as dissimilar. This showed that the funniness ratings were mostly influenced by the structure factor (16 of 29 stimuli), the offensiveness ratings were mostly influenced by the content factor (24 of 33 stimuli), and the boringness ratings were influenced by both the contents (8 of 24 stimuli) and structures (9 of 24 stimuli).

Validity of the Bimodal Model

The regressions were conducted separately for the three ratings, with the four content and structure factors as predictors and the 3 WD scores and gender as criteria (see Table 2). As expected, convergent validity was established between the funniness ratings of the INC-RES, NON, and SEX categories of the 3 WD with the corresponding funniness ratings of the three latent factors. The SEX-factor of the boringness ratings positively predicted the SEX category of the 3 WD (aversiveness rating; $p=.012$), and the INC-RES-factor significantly and negatively predicted the INC-RES-category of the 3 WD (funniness rating). The offensiveness ratings showed a convergence of the SEX-factor with the aversiveness ratings of the SEX-category in the 3 WD. The AGG-factor was spread across the three 3 WD categories and did not significantly predict any category. This is to be expected as aggressive contents were not targeted in the 3 WD.

In terms of gender, men found the SEX stimuli funnier and less boring than women, which is in line with the expectations. No significant prediction was found for the AGG stimuli. Finally, women found NON-stimuli more offensive than men, which was not hypothesized.

Table 2
Standardized Regression Weights with the Four Latent Factors as Predictors and Gender and the 3 WD Categories as Criteria

Rating & factor	Gender	3 WD funniness			3 WD aversiveness		
		INC-RES	NON	SEX	INC-RES	NON	SEX
Funniness							
INC-RES	.06	.65**	.21**	.37**	-.07	.09	.02
NON	.01	.06	.52**	.25**	.06	-.10	-.10
SEX	-.19**	.29**	.20	.65**	-.02	.10	-.42**
AGG	.08	.10	.22	-.25	-.18	-.27	.25
Offensiveness							
INC-RES	.03	.04	.25	.01	.03	-.11	.07
NON	.16**	.00	-.15	-.12	.08	.27	.36**
SEX	-.05	-.03	-.20	-.20	.24	.35**	.44**
AGG	.14	.16	.16	.18	-.02	-.01	-.03
Boringness							
INC-RES	.02	-.56**	.25	.01	.12	.03	-.06
NON	-.05	.05	-.15	-.12	.07	.25	-.18
SEX	.21**	-.02	-.20	-.20	.06	-.04	.49
AGG	-.11	-.36	.26	.18	-.02	.06	-.24

Note. $N=225$. NON=nonsense, INC-RES=incongruity-resolution, AGG=aggressive. Gender coded as 1=men, 2=women. Fully standardized solution reported for the 3 WD, while only latent variables were standardized for gender. Convergent regression weights in bold.

** $p<.01$.

Discussion

The first aim on the present study was to empirically test a bimodal model of humor appreciation (Ruch & Hehl, 2007; Ruch & Platt, 2012), in which two structure and content factors were specified. As expected, this model showed a descriptively better fit to alternative models. In other words, considering the structure and content of humorous stimuli simultaneously as latent factors better represented individual differences in funniness, offensiveness, and boringness ratings of a diverse set of jokes and cartoons. This is in line with previous theoretical and empirical work that took both stimulus properties into account (e.g., Carretero-Dios et al., 2010; Freud, 1905; Godkewitsch, 1976; Ruch, 1992). Neglecting

either contents or structures could lead to erroneous conclusions; for example, if humorous stimuli have different contents and structures, but only one of these properties is modeled, the results and interpretations would be biased (see also Ruch & Hehl, 2007; Ruch & Platt, 2012).

Additionally, the contribution of contents and structures to the appreciation of humorous stimuli differed across the individual stimuli and across the three ratings (funniness, offensiveness, and boringness). The bimodal approach used in the present study allows, for the first time, to adequately model these different sources of humor appreciation. This finding also supports the need for a careful stimuli selection and evaluation to specify the loadings in the bimodal models in future research. This should include expert ratings on the content and structure of the humorous stimuli to ensure their content validity (see Carretero-Dios, Pérez, & Buéla-Casal, 2009). In the present study, the stimuli were selected to have both a salient content and structure. Future studies should extend these ratings to further contents (e.g., dark, innocent/harmless, benevolent, or satirical) and structures (e.g., different kinds of incongruity-resolution or nonsense; see Raskin, 1985; Raskin, Hempelmann, & Taylor, 2009). This would allow testing how many content and structure factors can be distinguished, thus providing an empirical means of developing a comprehensive taxonomy of humor appreciation. Furthermore, if stimuli are included that mainly contain a salient content or structure, the bimodal model can also include latent correlations between the content and structure factors to better understand their mutual interplay.

The second aim of the present study was to investigate the construct and criterion validity of these two content and structure factors by relating them to the 3 WD and gender.

Convergent validity was established with the 3 WD categories for INC-RES, NON, and SEX for the funniness ratings. As anticipated, the convergence was lower for the offensiveness and boringness ratings, showing that the negative reactions to humorous stimuli (aversiveness vs. offensiveness and boringness) captured different aspects. Furthermore, the gender differences reported in the literature were replicated for SEX (men appreciating it more than women) and INC-RES (no gender differences), while the previous findings were not replicated for NON and AGG (Hofmann et al., 2019). As previous studies did however not simultaneously model the content and structure properties of the humorous stimuli, it is unclear to what extent these gender differences are comparable. Future studies should extend this initial validation by relating the content and structure factors to other individual-difference variables, such as self-reports of humor (e.g., styles related to mockery and nonsense; see Ruch et al., 2018).

Limitations and Future Directions

The present study only captured two structures and contents as a minimal test of the bimodal model of humor appreciation. Thus, replicating and extending the present approach to further structures and contents is an important next step in research. Comparing these latent factors shows how many active stimuli ingredients can be distinguished in humor appreciation. Also, investigating the nomological network of these contents and structures would showcase the relevance of humor appreciation in terms of traits, habits, abilities, attitudes, and interests, building on the idea that humor appreciation can serve as an implicit personality measure (e.g., Cattell & Tollefson, 1966).

Next, the stimuli in the present study were selected to capture both a salient content and structure. Selecting stimuli a-priori that primarily contain a certain structure or content as well as mixed stimuli would enable testing interactions between these two stimulus properties

in determining a person's evaluation of the funniness, offensiveness, and boringness of humorous stimuli (see Godkewitsch, 1976). Also, people likely differ in how they perceive and interpret the content and structure of a stimulus (i.e., individual differences in humor comprehension and the response dimensions of humor appreciation). This would provide another direction for future studies, for example, by having open responses on the comprehension of humorous stimuli rated for the different contents and structures, by using eye-tracking methods, or by having the stimuli rated for the extent to which different contents and structures are perceived (as was done by Wicker et al., 1981).

Furthermore, the large number of stimuli in the models resulted in a ratio of participants to estimated parameters of 2:1 to 3:1, rather than the recommended minimum of 10:1 (e.g., Kline, 2011). Thus, the obtained parameter estimates might not be stable and are in need of replication in new and larger samples. Importantly, additional analyses conducted using parcels (with a ratio of 6:1 to 9:1) replicated the present findings (three parcels for each of the 12 Content \times Structure \times Rating combinations using a radial parceling approach as recommended by Little, Rhemtulla, Gibson, & Schoemann, 2013), yet future studies should increase the ratio beyond 10:1. Also, post-hoc power analyses (conducted with the WebPower package, Zhang & Yuan, 2018; see Supplementary Tables S4–S6) showed that the power of 82.1% (133 of 162) of the estimated parameters was higher than .80. The code for the power analyses is provided in the supplementary materials to allow planning the sample sizes of future studies.

Lastly, females, younger adults, and well-educated people were overrepresented in the sample. Thus, future studies should employ larger and more representative samples to

replicate the present findings. Replications in different cultures (see e.g., Carretero-Dios & Ruch, 2010; Eysenck, 1944) would also be desirable.

Conclusion

Overall, the present study suggests that it is viable to model both contents and structures of humorous stimuli simultaneously to adequately represent individual differences in humor appreciation (Ruch & Hehl, 2007; Ruch & Platt, 2012). Such a bimodal model was found to outperform alternative models in all three ratings (funniness, offensiveness, and boringness). Furthermore, the construct and criterion validity of the content and structure factors in the model received initial support. This study thus showcases a promising future direction for theories, research, and applications of individual differences in humor appreciation, and possibly also other humor domains such as humor production and comprehension.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at <https://doi.org/10.1027/1614-0001/a000301>

ESM 1. This document provides detailed description of the initial stimuli selection, the empirical stimuli selection, additional analyses, and the code for the power analyses.

References

- Carretero-Dios, H., Pérez, C., & Buela-Casal, G. (2009). Content validity and metric properties of a pool of items developed to assess humor appreciation. *The Spanish Journal of Psychology*, 12, 773–787. <https://doi.org/10.1017/S1138741600002146>
- Carretero-Dios, H., Pérez, C., & Buela-Casal, G. (2010). Assessing the appreciation of the content and structure of humor: Construction of a new scale. *Humor: International Journal of Humor Research*, 23, 307–325. <https://doi.org/10.1515/humr.2010.014>.
- Carretero-Dios, H., & Ruch, W. (2010). Humor appreciation and sensation seeking: Invariance of findings across culture and assessment instrument? *Humor: International Journal of Humor Research*, 23, 427–445. <https://doi.org/10.1515/HUMR.2010.020>
- Cattell, R.B., & Tollefson, D.L. (1966). *IPAT humor test of personality*. Champaign, IL: Institute for Personality and Ability Testing.
- Eysenck, H.J. (1942). The appreciation of humour: An experimental and theoretical study. *British Journal of Psychology*, 32, 295–309. <https://doi.org/10.1111/j.2044-8295.1942.tb01027.x>
- Eysenck, H.J. (1944). National differences in “sense of humor”: Three experimental and statistical studies. *Journal of Personality*, 13, 37–54. <https://doi.org/10.1111/j.1467-6494.1944.tb01971.x>
- Ferguson, M.A., & Ford, T.E. (2008). Disparagement humor: A theoretical and empirical review of psychoanalytic, superiority, and social identity theories. *Humor: International Journal of Humor Research*, 21, 283–312. <https://doi.org/10.1515/HUMOR.2008.014>

Freud, S. (1905). *Der Witz und seine Beziehung zum Unbewussten*. [The joke and its relation to the unconscious.] Vienna, Austria: Deutike.

Godkewitsch, M. (1976). Thematic and collative properties of written jokes and their contribution to funniness. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 8, 88–97.

<https://doi.org/10.1037/h0081937>

Hofmann, J., Platt, T., Lau, C., & Torres-Marín, J. (2019). *Gender differences in humor traits, appreciation, production, comprehension, (neural) responses, use, and correlates: A systematic review*. Manuscript under review.

Kline, R.B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Little, T.D., Rhemtulla, M., Gibson, K., & Schoemann, A.M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, 18, 285–300.

<https://doi.org/10.1037/a0033266>

Martin, R. A., & Ford, T. E. (2018). *The psychology of humor: An integrative approach* (2nd ed.). Oxford, UK: Academic Press.

Mindess, H., Miller, C., Turek, J., Bender, A., & Corbin, S. (1985). *The Antioch humor test: Making sense of humor*. New York, NY: Avon.

Raskin, V. (1985). *Semantic mechanisms of humor*. Dordrecht, The Netherlands: D. Reidel Publishing Company.

Raskin, V., Hempelmann, C. F., & Taylor, J. M. (2009). How to understand and assess a theory: The evolution of the SSTH into the GTVH and now into the OSTH. *Journal of Literary Theory*, 3, 285–311. <https://doi.org/10.1515/JLT.2009.016>

- Reiss, S., & Reips, U.-D. (2016). Online assessment. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advances* (pp. 120–134). Göttingen, Germany: Hogrefe.
- Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, Illinois, USA. <https://CRAN.R-project.org/package=psych>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, <https://doi.org/10.18637/jss.v048.i02>
- Ruch, W. (1992). Assessment of appreciation of humor: Studies with the 3 WD Humor Test. In C.D. Spielberger & J.N. Butcher, *Advances in personality* (pp. 27–75). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ruch, W. (Ed.). (2007). *The sense of humor: Explorations of a personality characteristic* (2nd ed.). Berlin, Germany: Mouton de Gruyter.
- Ruch, W. (2012). Towards a new structural model of the sense of humor: Preliminary findings. In V. Raskin & J. M. Taylor (Chairs), *Artificial Intelligence of Humor: Papers from the 2012 AAAI Fall Symposium* (AAAI Technical Report FS-12-02, pp. 68-75). Retrieved from http://www.ilhaire.eu/pdf/Ruch_2012.pdf
- Ruch, W., & Hehl, F.-J. (2007). A two-mode model of humor appreciation: Its relation to aesthetic appreciation and simplicity-complexity of personality. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (2nd ed., pp. 109–142). Berlin, Germany: Mouton de Gruyter.

- Ruch, W., Heintz, S., Platt, T., Wagner, L., & Proyer, R. T. (2018). Broadening humor: Comic styles differentially tap into temperament, character, and ability. *Frontiers in Psychology*, 9: 6. <https://doi.org/10.3389/fpsyg.2018.00006>
- Ruch, W., & Hofmann, J. (2012). A temperament approach to humor. In P. Gremigni (Ed.), *Humor and health promotion* (pp. 79–113). Hauppauge, NY: Nova Science Publ.
- Ruch, W., & Platt, T. (2012). Separating content and structure in humor appreciation: The need for a bimodal model and support from research into aesthetics. In A. Nijhold (Ed.), *3rd International workshop on computational humor* (pp. 23–27). Amsterdam, The Netherlands: University of Amsterdam.
- Ruch, W., & Rath, S. (1993). The nature of humor appreciation: Toward an integration of perception of stimulus properties and affective experience. *Humor: International Journal of Humor Research*, 6, 363–384. <https://doi.org/10.1515/humr.1993.6.4.363>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74. Retrieved from <http://www.mpr-online.de>
- Suls, J. (1983). Cognitive processes in humor appreciation. In P.E. McGhee & J.H. Goldstein (Eds.), *Handbook of humor research* (pp. 39–57). New York, NY: Springer.
- Wicker, F.W., Thorelli, I.M., Barron, W.L., & Ponder, M.R. (1981). Relationships among affective and cognitive factors in humor. *Journal of Research in Personality*, 15, 359–370. [https://doi.org/10.1016/0092-6566\(81\)90033-7](https://doi.org/10.1016/0092-6566(81)90033-7)
- Zhang, Z., & Yuan, K.-H. (Eds.) (2018). *Practical statistical power analysis using Webpower and R*. Granger, IN: ISDSA Press.